DT01 Rec'd PCT/FTT 2 0 JAN 2005

SEQUENCE LISTING

```
<110> BASF Aktiengesellschaft
<120> Process for the biotransformation of carotenoids
<130> M43191 beta-Karotin Biotransformation
<140>
<141>
<160> 12
<170> PatentIn Ver. 2.1
<210> 1
<211> 1170
<212> DNA
<213> Thermus thermophilus
<220>
<221> CDS
<222> (1)..(1170)
 <400> 1
atg aag cgc ctt tcc ctg agg gag gcc tgg ccc tac ctg aaa gac ctc
Met Lys Arg Leu Ser Leu Arg Glu Ala Trp Pro Tyr Leu Lys Asp Leu
                                                           15
                                      10
   1
cag caa gat ccc ctc gcc gtc ctg ctg gcg tgg ggc cgg gcc cac ccc
                                                                    96
Gln Gln Asp Pro Leu Ala Val Leu Leu Ala Trp Gly Arg Ala His Pro
                                                       30
                                  25
              20
 cgg ctc ttc ctt ccc ctg ccc cgc ttc ccc ctg gcc ctg atc ttt gac
                                                                    144
Arg Leu Phe Leu Pro Leu Pro Arg Phe Pro Leu Ala Leu Ile Phe Asp
          35
 ccc gag ggg gtg gag ggg ctc ctc gcc gag ggg acc acc aag gcc
Pro Glu Gly Val Glu Gly Ala Leu Leu Ala Glu Gly Thr Thr Lys Ala
                                               60
                          55
      50
 acc ttc cag tac cgg gcc ctc tcc cgc ctc acg ggg agg ggc ctc ctc
                                                                    240
 Thr Phe Gln Tyr Arg Ala Leu Ser Arg Leu Thr Gly Arg Gly Leu Leu
                                                               80
                                           75
                      70
  65
 acc gac tgg ggg gaa agc tgg aag gag gcg cgc aag gcc ctc aaa gac
                                                                     288
 Thr Asp Trp Gly Glu Ser Trp Lys Glu Ala Arg Lys Ala Leu Lys Asp
                                                           95
                                       90
                  85
 ccc ttc ctg ccg aag aac gtc cgc ggc tac cgg gag gcc atg gag gag
                                                                     336
 Pro Phe Leu Pro Lys Asn Val Arg Gly Tyr Arg Glu Ala Met Glu Glu
                                                      110
                                  105
             100
 gag gcc cgg gcc ttc ttc ggg gag tgg cgg ggg gag gag cgg gac ctg
                                                                     384
 Glu Ala Arg Ala Phe Phe Gly Glu Trp Arg Gly Glu Glu Arg Asp Leu
                                                  125
                              120
         115
 gac cac gag atg ctc gcc ctc tcc ctg cgc ctc ctc ggg cgg gcc ctc
                                                                     432
 Asp His Glu Met Leu Ala Leu Ser Leu Arg Leu Leu Gly Arg Ala Leu
                                              140
                          135
     130
 ttc ggg aag ccc ctc tcc cca agc ctc gcg gag cac gcc ctt aag gcc
                                                                     480
 Phe Gly Lys Pro Leu Ser Pro Ser Leu Ala Glu His Ala Leu Lys Ala
```

.

	145					150					155					160	
	ctg (gac Asp	cgg Arg	atc Ile	atg Met 165	gcc Ala	cag Gln	acc . Thr	Arg	agc Ser 170	ccc Pro	ctg Leu	gcc (Ala :	LVu .	ctg Leu 175	gac Asp	528
	ctg Leu	gcc Ala	gcc Ala	gaa Glu 180	gcc Ala	cgc Arg	ttc Phe	Arg	aag Lys 185	gac Asp	cgg Arg	ggg Gly	AIG	ctc Leu 190	tac Tyr	cgc Arg	576
-	gag Glu	gcg Ala	gaa Glu 195	gcc Ala	ctc Leu	atc Ile	gtc Val	cac His 200	ccg Pro	ccc Pro	ctc Leu	tcc Ser	cac His 205	ctt Leu	ccc Pro	cga Arg	624
•	gag Glu	cgc Arg 210	gcc Ala	ctg Leu	agc Ser	gag Glu	gcc Ala 215	gtg Val	acc Thr	ctc Leu	ctg Leu	gtg Val 220	gcg Ala	ggc Gly	cac His	gag Glu	672
	acg Thr 225	gtg Val	gcg Ala	agc Ser	gcc Ala	ctc Leu 230	acc Thr	tgg Trp	tcc Ser	ttt Phe	ctc Leu 235	ctc Leu	ctc Leu	tcc Ser	cac His	cgc Arg 240	720
	ccg Pro	gac Asp	tgg Trp	cag Gln	aag Lys 245	cgg Arg	gtg Val	gcc Ala	gag Glu	agc Ser 250	GIU	gag Glu	gcg Ala	gcc Ala	ctc Leu 255	gcc Ala	768
	gcc Ala	ttc Phe	cag Gln	gag Glu 260	Ala	ctg Leu	agg Arg	ctc Leu	tac Tyr 265	PIO	ccc Pro	gcc Ala	tgg Trp	atc Ile 270	ctc Leu	acc Thr	816
	cgg Arg	agg Arg	cto Leu 275	ı Glu	agg Arg	ccc	ctc Leu	ctc Leu 280	Leu	gga Gly	gag Glu	gac Asp	cgg Arg 285	ctc Leu	ccc Pro	ccg Pro	864
	ggc Gly	aco Thr	Thi	c ctc Lei	g gto ı Val	ctc Leu	tcc Ser 295	Pro	tac Tyr	gto Val	acc Thr	cag Gln 300	, ALS	ctc Leu	cac	ttc Phe	912
	ccc Pro	As _I	gg Gly	g gaq y Gli	g gco ı Ala	ttc Phe	Arg	ccc Pro	gag Glu	g cgc	tto Phe		g gag ı Glu	gaa Glu	agg Arg	999 Gly 320	960
	acc Thr	c cc	t to	g ggg	g cgo y Arg 325	g Tyr	tto Phe	ccc Pro	ttt Phe	gg(Gl ₃ 33(у пес	g ggg	g cag / Gln	agg Arg	cto Lev 335	tgc Cys	1008
	cto Lei	1 Gl	g cg y Ar	g ga g As 34	p Pho	c gco e Ala	cto Lei	cto Lei	gag ı Glu 34!	r GT	c cco	c ato	gto Val	cto Leu 350		g gcc g Ala	1056
	tto Phe	c tt e Ph	c cg e Ar 35	g Ar	c tte g Ph	c cgo	c cta g Lei	a gad 1 Asp 360	D PIG	c cto	c cce u Pre	c tto o Pho	c ccc e Pro 365		g gto g Val	c ctc l Leu	1104
	gc: Ala	c ca a Gl 37	n Va	c ac l Th	c ct r Le	g agg u Arg	g cco g Pro 37!	o GII	a gg u Gl	y Gl	g ct	t cc u Pro 38	·	g cgg a Arg	g cct	t agg	1152
	ga Gl 38	u Gl	g gt u Va	g cg l Ar	g gc g Al	g tga a 39											1170

- <210> 2
- <211> 389
- <212> PRT
- <213> Thermus thermophilus
- <400> 2
- Met Lys Arg Leu Ser Leu Arg Glu Ala Trp Pro Tyr Leu Lys Asp Leu
- Gln Gln Asp Pro Leu Ala Val Leu Leu Ala Trp Gly Arg Ala His Pro
- Arg Leu Phe Leu Pro Leu Pro Arg Phe Pro Leu Ala Leu Ile Phe Asp
- Pro Glu Gly Val Glu Gly Ala Leu Leu Ala Glu Gly Thr Thr Lys Ala
- Thr Phe Gln Tyr Arg Ala Leu Ser Arg Leu Thr Gly Arg Gly Leu Leu
- Thr Asp Trp Gly Glu Ser Trp Lys Glu Ala Arg Lys Ala Leu Lys Asp
- Pro Phe Leu Pro Lys Asn Val Arg Gly Tyr Arg Glu Ala Met Glu Glu
- Glu Ala Arg Ala Phe Phe Gly Glu Trp Arg Gly Glu Glu Arg Asp Leu
- Asp His Glu Met Leu Ala Leu Ser Leu Arg Leu Leu Gly Arg Ala Leu
- Phe Gly Lys Pro Leu Ser Pro Ser Leu Ala Glu His Ala Leu Lys Ala
- Leu Asp Arg Ile Met Ala Gln Thr Arg Ser Pro Leu Ala Leu Leu Asp
- Leu Ala Ala Glu Ala Arg Phe Arg Lys Asp Arg Gly Ala Leu Tyr Arg
- Glu Ala Glu Ala Leu Ile Val His Pro Pro Leu Ser His Leu Pro Arg
- Glu Arg Ala Leu Ser Glu Ala Val Thr Leu Leu Val Ala Gly His Glu
- Thr Val Ala Ser Ala Leu Thr Trp Ser Phe Leu Leu Leu Ser His Arg
- Pro Asp Trp Gln Lys Arg Val Ala Glu Ser Glu Glu Ala Ala Leu Ala
- Ala Phe Gln Glu Ala Leu Arg Leu Tyr Pro Pro Ala Trp Ile Leu Thr
- Arg Arg Leu Glu Arg Pro Leu Leu Leu Gly Glu Asp Arg Leu Pro Pro
- Gly Thr Thr Leu Val Leu Ser Pro Tyr Val Thr Gln Arg Leu His Phe

```
Pro Asp Gly Glu Ala Phe Arg Pro Glu Arg Phe Leu Glu Glu Arg Gly
                                                             320
                                         315
                    310
305
Thr Pro Ser Gly Arg Tyr Phe Pro Phe Gly Leu Gly Gln Arg Leu Cys
                                                         335
                                     330
                325
Leu Gly Arg Asp Phe Ala Leu Leu Glu Gly Pro Ile Val Leu Arg Ala
                                                     350
                                 345
            340
Phe Phe Arg Arg Phe Arg Leu Asp Pro Leu Pro Phe Pro Arg Val Leu
                                                 365
                             360
        355
Ala Gln Val Thr Leu Arg Pro Glu Gly Gly Leu Pro Ala Arg Pro Arg
                                             380
                         375
    370
Glu Glu Val Arg Ala
385
<210> 3
<211> 1188
<212> DNA
<213> Artificial sequence
<220>
<221> misc_feature
<222> (4)..(21)
<223> His tag
<220>
<223> Description of the artificial sequence: N-terminal
      his tagged
<220>
 <221> CDS
<222> (1)..(1188)
<400> 3
atg cat cac cat cat cac aag cgc ctt tcc ctg agg gag gcc tgg
Met His His His His His Lys Arg Leu Ser Leu Arg Glu Ala Trp
                                                           15
                                       10
                   5
   1
ccc tac ctg aaa gac ctc cag caa gat ccc ctc gcc gtc ctg ctg gcg
                                                                    96
Pro Tyr Leu Lys Asp Leu Gln Gln Asp Pro Leu Ala Val Leu Leu Ala
                                                       30
                                   25
              20
 tgg ggc cgg gcc cac ccc cgg ctc ttc ctt ccc ctg ccc cgc ttc ccc
                                                                    144
 Trp Gly Arg Ala His Pro Arg Leu Phe Leu Pro Leu Pro Arg Phe Pro
                                                   45
                               40
          35
 ctg gcc ctg atc ttt gac ccc gag ggg gtg gag ggg gcg ctc ctc gcc
                                                                    192
Leu Ala Leu Ile Phe Asp Pro Glu Gly Val Glu Gly Ala Leu Leu Ala
                          55
      50
 gag ggg acc acc aag gcc acc ttc cag tac cgg gcc ctc tcc cgc ctc
                                                                     240
 Glu Gly Thr Thr Lys Ala Thr Phe Gln Tyr Arg Ala Leu Ser Arg Leu
                                                               80
                                           75
                      70
  65
 acg ggg agg ggc ctc ctc acc gac tgg ggg gaa agc tgg aag gag gcg
                                                                     288
 Thr Gly Arg Gly Leu Leu Thr Asp Trp Gly Glu Ser Trp Lys Glu Ala
```

90

85

95

cgc Arg	aag Lys	gcc Ala	ctc Leu 100	aaa Lys	gac Asp	ccc Pro	ttc Phe	ctg Leu 105	ccg Pro	aag Lys	aac Asn	gtc Val	cgc Arg 110	ggc	tac Tyr	336
cgg Arg	gag Glu	gcc Ala 115	atg Met	gag Glu	gag Glu	gag Glu	gcc Ala 120	cgg Arg	gcc Ala	ttc Phe	ttc Phe	999 Gly 125	gag Glu	tgg Trp	cgg Arg	384
Gly ggg	gag Glu 130	gag Glu	cgg Arg	gac Asp	ctg Leu	gac Asp 135	cac His	gag Glu	atg Met	ctc Leu	gcc Ala 140	ctc Leu	tcc Ser	ctg Leu	cgc Arg	432
ctc Leu 145	ctc Leu	Gly 999	cgg Arg	gcc Ala	ctc Leu 150	ttc Phe	ggg Gly	aag Lys	ccc Pro	ctc Leu 155	tcc Ser	cca Pro	agc Ser	ctc Leu	gcg Ala 160	480
gag Glu	cac	gcc Ala	ctt Leu	aag Lys 165	gcc Ala	ctg Leu	gac Asp	cgg Arg	atc Ile 170	atg Met	gcc Ala	cag Gln	acc Thr	agg Arg 175	agc Ser	528
ccc Pro	ctg Leu	gcc Ala	ctc Leu 180	ctg Leu	gac Asp	ctg Leu	gcc Ala	gcc Ala 185	gaa Glu	gcc Ala	cgc Arg	ttc Phe	cgg Arg 190	aag Lys	gac Asp	576
cgg Arg	gly ggg	gcc Ala 195	ctc Leu	tac Tyr	cgc Arg	gag Glu	gcg Ala 200	gaa Glu	gcc Ala	ctc Leu	atc Ile	gtc Val 205	cac His	ccg Pro		624
ctc Leu	tcc Ser 210	cac His	ctt Leu	ccc Pro	cga Arg	gag Glu 215	cgc Arg	gcc Ala	ctg Leu	agc Ser	gag Glu 220	Ala	gtg Val	acc Thr		672
ctg Leu 225	Val	gcg Ala	ggc Gly	cac His	gag Glu 230	acg Thr	gtg Val	gcg Ala	agc Ser	gcc Ala 235	ctc Leu	acc Thr	tgg Trp	tcc Ser	ttt Phe 240	720
ctc Leu	ctc Leu	ctc Leu	tcc Ser	cac His 245	Arg	ccg Pro	gac Asp	tgg Trp	cag Gln 250	Lys	cgg Arg	gtg Val	gcc Ala	gag Glu 255	agc Ser	768
gag Glu	gag Glu	gcg Ala	gcc Ala 260	Leu	gcc Ala	gcc Ala	ttc Phe	cag Gln 265	gag Glu	gcc Ala	ctg Leu	agg Arg	ctc Leu 270	Tyr	ccc Pro	816
ccc Pro	gcc Ala	tgg Trp 275	Ile	ctc Leu	acc Thr	cgg Arg	agg Arg 280	Leu	gaa Glu	agg Arg	ccc Pro	ctc Leu 285	Leu	ctg Leu	gga Gly	864
gag Glu	gac Asp 290	Arg	ctc Leu	ccc Pro	ccg Pro	ggc Gly 295	Thr	acc Thr	ctg Leu	gtc Val	ctc Leu 300	Ser	ccc Pro	tac Tyr	gtg Val	912
acc Thr 305	Gln	agg Arg	ctc Leu	cac His	ttc Phe 310	Pro	gat Asp	ggg Gly	gag Glu	gcc Ala 315	Phe	cgg Arg	cco Pro	gag Glu	cgc Arg 320	960
ttc Phe	ctg Leu	gag Glu	gaa Glu	agg Arg 325	Gly	acc Thr	cct Pro	tcg Ser	ggg 330	Arg	tac Tyr	ttc Phe	ccc Pro	Phe 335	ggc Gly	1008
													cto			1056

ccc atc gtc ctc agg gcc ttc ttc cgc cgc ttc cgc cta gac ccc ctc Pro Ile Val Leu Arg Ala Phe Phe Arg Arg Phe Arg Leu Asp Pro Leu ccc ttc ccc cgg gtc ctc gcc cag gtc acc ctg agg ccc gaa ggc ggg Pro Phe Pro Arg Val Leu Ala Gln Val Thr Leu Arg Pro Glu Gly Gly ctt ccc gcg cgg cct agg gag gag gtg cgg gcg tga Leu Pro Ala Arg Pro Arg Glu Glu Val Arg Ala <210> 4 <211> 395 <212> PRT <213> Artificial sequence <223> Description of the artificial sequence: N-terminal his tagged <400> 4 Met His His His His His Lys Arg Leu Ser Leu Arg Glu Ala Trp Pro Tyr Leu Lys Asp Leu Gln Gln Asp Pro Leu Ala Val Leu Leu Ala Trp Gly Arg Ala His Pro Arg Leu Phe Leu Pro Leu Pro Arg Phe Pro Leu Ala Leu Ile Phe Asp Pro Glu Gly Val Glu Gly Ala Leu Leu Ala Glu Gly Thr Thr Lys Ala Thr Phe Gln Tyr Arg Ala Leu Ser Arg Leu Thr Gly Arg Gly Leu Leu Thr Asp Trp Gly Glu Ser Trp Lys Glu Ala Arg Lys Ala Leu Lys Asp Pro Phe Leu Pro Lys Asn Val Arg Gly Tyr Arg Glu Ala Met Glu Glu Glu Ala Arg Ala Phe Phe Gly Glu Trp Arg Gly Glu Glu Arg Asp Leu Asp His Glu Met Leu Ala Leu Ser Leu Arg Leu Leu Gly Arg Ala Leu Phe Gly Lys Pro Leu Ser Pro Ser Leu Ala Glu His Ala Leu Lys Ala Leu Asp Arg Ile Met Ala Gln Thr Arg Ser Pro Leu Ala Leu Leu Asp Leu Ala Ala Glu Ala Arg Phe Arg Lys Asp Arg Gly Ala Leu Tyr Arg Glu Ala Glu Ala Leu Ile Val His Pro Pro Leu Ser His Leu Pro Arg Glu Arg Ala Leu Ser Glu Ala Val Thr Leu

Leu Val Ala Gly His Glu Thr Val Ala Ser Ala Leu Thr Trp Ser Phe 240 235 230 225 Leu Leu Ser His Arg Pro Asp Trp Gln Lys Arg Val Ala Glu Ser 255 250 245 Glu Glu Ala Ala Leu Ala Ala Phe Gln Glu Ala Leu Arg Leu Tyr Pro 270 265 260 Pro Ala Trp Ile Leu Thr Arg Arg Leu Glu Arg Pro Leu Leu Gly 285 280 275 Glu Asp Arg Leu Pro Pro Gly Thr Thr Leu Val Leu Ser Pro Tyr Val 300 295 290 Thr Gln Arg Leu His Phe Pro Asp Gly Glu Ala Phe Arg Pro Glu Arg 320 315 310 305 Phe Leu Glu Glu Arg Gly Thr Pro Ser Gly Arg Tyr Phe Pro Phe Gly 335 330 325 Leu Gly Gln Arg Leu Cys Leu Gly Arg Asp Phe Ala Leu Leu Glu Gly 350 345 340 Pro Ile Val Leu Arg Ala Phe Phe Arg Arg Phe Arg Leu Asp Pro Leu 365 360 355 Pro Phe Pro Arg Val Leu Ala Gln Val Thr Leu Arg Pro Glu Gly Gly 380 375 370 Leu Pro Ala Arg Pro Arg Glu Glu Val Arg Ala 395 390 385 <210> 5 <211> 1188 <212> DNA <213> Artificial sequence <220> <221> misc_feature <222> (1168)..(1185) <223> His tag <220> <223> Description of the artificial sequence: N-terminal His-tagged <220> <221> CDS <222> (1)..(1188) <400> 5 atg aag cgc ctt tcc ctg agg gag gcc tgg ccc tac ctg aaa gac ctc 48 Met Lys Arg Leu Ser Leu Arg Glu Ala Trp Pro Tyr Leu Lys Asp Leu 15 10 5 1 cag caa gat ccc ctc gcc gtc ctg ctg gcg tgg ggc cgg gcc cac ccc 96

Gln Gln Asp Pro Leu Ala Val Leu Leu Ala Trp Gly Arg Ala His Pro

20

25

30

cgg Arg	ctc Leu	ttc Phe 35	ctt Leu	ccc Pro	ctg Leu	ccc Pro	cgc Arg 40	ttc Phe	ccc Pro	ctg Leu	gcc Ala	ctg Leu 45	atc Ile	ttt Phe	gac Asp	144
ccc Pro	gag Glu 50	gly ggg	gtg Val	gag Glu	G1y 999	gcg Ala 55	ctc Leu	ctc Leu	gcc Ala	gag Glu	999 Gly 60	acc Thr	acc Thr	aag Lys	gcc Ala	192
acc Thr 65	ttc Phe	cag Gln	tac Tyr	cgg Arg	gcc Ala 70	ctc Leu	tcc Ser	cgc Arg	ctc Leu	acg Thr 75	Gly 999	agg Arg	ggc Gly	ctc Leu	ctc Leu 80	240
acc Thr	gac Asp	tgg Trp	gly ggg	gaa Glu 85	agc Ser	tgg Trp	aag Lys	gag Glu	gcg Ala 90	cgc Arg	aag Lys	gcc Ala	ctc Leu	aaa Lys 95	gac Asp	288
ccc Pro	ttc Phe	ctg Leu	ccg Pro 100	aag Lys	aac Asn	gtc Val	cgc Arg	ggc Gly 105	tac Tyr	cgg Arg	gag Glu	gcc Ala	atg Met 110	gag Glu	gag Glu	336
gag Glu	gcc Ala	cgg Arg 115	gcc Ala	ttc Phe	ttc Phe	Gly aaa	gag Glu 120	tgg Trp	cgg Arg	gly	gag Glu	gag Glu 125	cgg Arg	gac Asp	ctg Leu	384
gac Asp	cac His 130	gag Glu	atg Met	ctc Leu	gcc Ala	ctc Leu 135	tcc Ser	ctg Leu	cgc Arg	ctc Leu	ctc Leu 140	Gly 999	cgg Arg	gcc Ala	ctc Leu	432
ttc Phe 145	Gly	aag Lys	ccc Pro	ctc Leu	tcc Ser 150	cca Pro	agc Ser	ctc Leu	gcg Ala	gag Glu 155	HIS	gcc Ala	ctt Leu	aag Lys	gcc Ala 160	480
ctg Leu	gac Asp	cgg Arg	atc Ile	atg Met 165	Ala	cag Gln	acc Thr	agg Arg	agc Ser 170	Pro	ctg Leu	gcc Ala	ctc Leu	ctg Leu 175	gac Asp	528
ctg Leu	gcc Ala	gcc Ala	gaa Glu 180	Ala	cgc Arg	ttc Phe	cgg Arg	aag Lys 185	Asp	cgg Arg	gly Gly	gcc Ala	ctc Leu 190	тут	cgc Arg	576
gag Glu	gcg Ala	gaa Glu 195	Ala	ctc Leu	atc Ile	gtc Val	cac His 200	Pro	ccc Pro	ctc Leu	tcc Ser	cac His 205	ьeu	ccc Pro	cga Arg	624
gag	g cgc l Arg 210	Ala	ctg Lev	g ago	gag Glu	gcc Ala 215	Val	acc Thr	ctc Leu	ctg Leu	g gtg Val 220	Ата	ggo	cac His	gag Glu	672
acc Thr 225	val:	gcg Ala	g ago	gcc Ala	ctc Leu 230	Thr	tgg Trp	tcc Ser	ttt Phe	cto Lev 235	ı ren	ctc Leu	tcc Ser	cac His	cgc Arg 240	720
ccs Pro	g gad o Asp	tgg Trp	g cag Glr	g aag n Lys 245	Arg	gtg Val	gcc Ala	gag Glu	ago Ser 250	GIU	g gag ı Glu	gcg Ala	gcc Ala	cto Lev 255	gcc Ala	768
gco	tto a Phe	c cag	g gag n Glu 260	ı Ala	ctg Leu	agg Arg	g cto g Lev	tac Tyr 265	Pro	cco Pro	gco Ala	tgg Trp	ato 11e 270	i net	acc Thr	816
cgg Arg	g agg	g ctg g Lei 275	ı Glu	a agg	Pro) Let	ı Lev	ı Lev	gga Gly	/ GI	ı Asp	c cgg Arg 285	nec	c cco	c ccg Pro	864

ggc Gly	acc Thr 290	acc Thr	ctg Leu	gtc Val	ctc Leu	tcc Ser 295	ccc Pro	tac Tyr	gtg Val	acc Thr	cag Gln 300	agg Arg	ctc Leu	cac His	ttc Phe	912
ccc Pro 305	gat Asp	ggg Gly	gag Glu	gcc Ala	ttc Phe 310	cgg Arg	ccc Pro	gag Glu	cgc Arg	ttc Phe 315	ctg Leu	gag Glu	gaa Glu	agg Arg	999 Gly 320	960
acc Thr	cct Pro	tcg Ser	ggg ggg	cgc Arg 325	tac Tyr	ttc Phe	ccc Pro	ttt Phe	ggc Gly 330	ctg Leu	gly ggg	cag Gln	agg Arg	ctc Leu 335	tgc Cys	1008
ctg Leu	Gly 999	cgg Arg	gac Asp 340	ttc Phe	gcc Ala	ctc Leu	ctc Leu	gag Glu 345	ggc Gly	ccc Pro	atc Ile	gtc Val	ctc Leu 350	agg Arg	gcc Ala	1056
ttc Phe	ttc Phe	cgc Arg 355	cgc Arg	ttc Phe	cgc Arg	cta Leu	gac Asp 360	ccc Pro	ctc Leu	ccc Pro	ttc Phe	ccc Pro 365	cgg Arg	gtc Val	ctc Leu	1104
gcc Ala	cag Gln 370	gtc Val	acc Thr	ctg Leu	agg Arg	ccc Pro 375	gaa Glu	ggc Gly	ggg Gly	ctt Leu	ccc Pro 380	gcg Ala	cgg Arg	cct Pro	agg Arg	1152
gag Glu 385	gag Glu	gtg Val	cgg Arg	gcg Ala	cat His 390	cac His	cat His	cat His	cat His	cac His 395	tga					1188
<21 <21 <22		RT rtif escr	icia iptic agge	on o	quen f th	ce e ar	tifi	cial	seq	uenc	e: C	-ter	mina	1		
<21 <21 <22	2 > P1 3 > A: 3 > D: H: 0 > 6 Lys	RT rtif escr is-t	iptio aggeo	on o: d	f th	e ar				Pro					Leu	
<21 <21 <22 <40 Met	2 > P1 3 > A: 3 > D: H: 0 > 6 Lys	RT rtif escr is-t Arg	iptio aggeo Leu	on of d Ser 5	f the	e ar	Glu	Ala Leu	Trp 10	Pro Trp	Tyr Gly	Leu Arg	Lys	Asp 15 His		
<21 <21 <22 <40 Met 1	2 > P1 3 > A: 3 > D: H: 0 > 6 Lys	RT rtif escr is-t Arg	ipticagged Leu Pro 20 Leu	on o	Leu Ala	Arg	Glu Leu	Ala Leu 25	Trp 10 Ala	Pro	Tyr	Leu	Lys Ala 30	Asp 15 His	Pro	
<21 <22 <40 Met 1 Gln	2 > Pl 3 > A: 3 > Do H: 0 > 6 Lys	RT rtif escr is-t Arg Asp Phe 35 Gly	ipticagged Leu Pro 20 Leu	on of d Ser 5 Leu Pro	Leu Ala Leu	e ar	Glu Leu Arg 40 Leu	Ala Leu 25 Phe	Trp 10 Ala Pro	Pro Trp	Tyr Gly Ala	Leu Arg Leu 45	Lys Ala 30 Ile	Asp 15 His	Pro Asp	
<21 <21 <22 <40 Met Gln Arg	2> Pl 3> A: 3> Do H 0> 6 Lys Gln Glu 50	RT rtif escr is-t Arg Asp Phe 35 Gly	ipticagged Leu Pro 20 Leu Val	on o	Leu Ala Leu Gly	e ar	Glu Leu Arg 40 Leu	Ala Leu 25 Phe	Trp 10 Ala Pro	Pro Trp Leu Glu	Tyr Gly Ala Gly 60	Leu Arg Leu 45 Thr	Lys Ala 30 Ile	Asp 15 His Phe	Pro Asp Ala	
<21 <21 <22 <40 Met Gln Arg Pro	2> Pl 3> A: 3> Do H 0> 6 Lys Gln Glu 50	RT rtif escr is-t Arg Asp Phe 35 Gly	Leu Pro 20 Leu Val	on of d Ser 5 Leu Pro Glu Arg	Leu Ala Gly Ala 70 Ser	e ar	Glu Leu Arg 40 Leu	Ala Leu 25 Phe Leu	Trp 10 Ala Pro Ala Leu	Pro Trp Leu Glu Thr 75	Tyr Gly Ala Gly 60	Leu Arg Leu 45 Thr	Lys Ala 30 Ile Thr	Asp 15 His Phe Lys	Pro Asp Ala Leu 80	
<21 <21 <22 <40 Met Gln Arg Pro Thr 65	2> Pl 3> A: 3> Do H 0> 6 Lys Gln Glu 50 Phe	RT rtif escr is-t Arg Asp Phe 35 Gly Gln Trp	Leu Pro 20 Leu Val Gly	on or	Leu Ala Gly Ala 70 Ser	Arg Val Pro Ala 55 Leu Trp	Glu Leu Arg 40 Leu Ser	Ala Leu 25 Phe Leu Arg	Trp 10 Ala Pro Ala Leu Ala 90	Pro Trp Leu Glu Thr 75 Arg	Tyr Gly Ala Gly 60 Gly	Leu Arg Leu 45 Thr	Lys Ala 30 Ile Thr Gly	Asp 15 His Phe Lys 95 Glu	Pro Asp Ala Leu 80 Asp	
<21 <21 <22 <40 Met Gln Arg Pro Thr 65	2> Pl 3> A: 3> Do H 0> 6 Lys Gln Glu 50 Phe	RT rtif escr is-t Arg Asp Phe 35 Gly Gln Trp Leu	Leu Pro 20 Leu Val Tyr Gly Pro 100 Ala	on or d Ser 5 Leu Pro Glu Arg Glu 85 Lys	Leu Ala Gly Ala 70 Ser	Arg Val Ala 55 Leu Trp	Glu Leu Arg 40 Leu Ser Lys	Ala Leu 25 Phe Leu Arg Glu Gly 105	Trp 10 Ala Pro Ala Leu Ala 90 Tyr	Pro Trp Leu Glu Thr 75 Arg	Tyr Gly Ala Gly 60 Gly Lys	Leu Arg Leu 45 Thr Arg	Lys Ala 30 Ile Thr Gly Leu Met 110 Arg	Asp 15 His Phe Lys 95 Glu	Pro Asp Ala Leu 80 Asp Glu	

130 135 140

Phe Gly Lys Pro Leu Ser Pro Ser Leu Ala Glu His Ala Leu Lys Ala 145 150 150

Leu Asp Arg Ile Met Ala Gln Thr Arg Ser Pro Leu Ala Leu Leu Asp 165 170 175

Leu Ala Ala Glu Ala Arg Phe Arg Lys Asp Arg Gly Ala Leu Tyr Arg 180 185 190

. Glu Ala Glu Ala Leu Ile Val His Pro Pro Leu Ser His Leu Pro Arg 195 200 205

- Glu Arg Ala Leu Ser Glu Ala Val Thr Leu Leu Val Ala Gly His Glu 210 220

Thr Val Ala Ser Ala Leu Thr Trp Ser Phe Leu Leu Leu Ser His Arg 225 230 230 235

Pro Asp Trp Gln Lys Arg Val Ala Glu Ser Glu Glu Ala Ala Leu Ala 245 250 255

Ala Phe Gln Glu Ala Leu Arg Leu Tyr Pro Pro Ala Trp Ile Leu Thr 260 265 270

Arg Arg Leu Glu Arg Pro Leu Leu Leu Gly Glu Asp Arg Leu Pro Pro 275 280 285

Gly Thr Thr Leu Val Leu Ser Pro Tyr Val Thr Gln Arg Leu His Phe 290 295 300

Pro Asp Gly Glu Ala Phe Arg Pro Glu Arg Phe Leu Glu Glu Arg Gly 305 310 315

Thr Pro Ser Gly Arg Tyr Phe Pro Phe Gly Leu Gly Gln Arg Leu Cys 325

Leu Gly Arg Asp Phe Ala Leu Leu Glu Gly Pro Ile Val Leu Arg Ala 340 345 350

Phe Phe Arg Arg Phe Arg Leu Asp Pro Leu Pro Phe Pro Arg Val Leu 355

Ala Gln Val Thr Leu Arg Pro Glu Gly Gly Leu Pro Ala Arg Pro Arg 370 375 380

Glu Glu Val Arg Ala His His His His His 395

<210> 7

<211> 30 <212> DNA

<213> Artificial sequence

<220>

<223> Description of the artificial sequence: PCR primer

<400> 7 cgaagctcat atgaagcgcc tttccctgag

```
<210> 8
 <211> 30
 <212> DNA
 <213> Artificial sequence
 <220>
 <223> Description of the artificial sequence: PCR primer
 <400> 8
                                                                     30
 gcgaattcac gcccgcacct cctccctagg
 <210> 9
<211> 42
 <212> DNA
 <213> Artificial sequence
  <220>
  <223> Description of the artificial sequence: PCR primer
  <400> 9
                                                                     42
  cgaagctcat atgcatcacc atcatcatca caagcgcctt tc
  <210> 10
  <211> 42
  <212> DNA
  <213> Artificial sequence
  <220>
  <223> Description of the artificial sequence: PCR primer
  <400> 10
                                                                      42
  cggaattcag tgatgatgat ggtgatgcgc ccgcacctcc tc
  <210> 11
  <211> 30
  <212> DNA
  <213> Artificial sequence
  <220>
  <223> Description of the artificial sequence: PCR primer
  <400> 11
                                                                      30
  ccggaattca tgaagcgcct ttccctgagg
  <210> 12
  <211> 44
  <212> DNA
   <213> Artificial sequence
   <220>
   <223> Description of the artificial sequence: PCR primer
   <400> 12
                                                                      44
   ccaatgcatt ggttctgcag tcaggcccgc acctcctccc tagg
```